

10266

10266

RESULT 8
 ATHRIENUCP
 LOCUS 1160 bp mRNA linear PLN 27-APR-1993
 DEFINITION Arabidopsis thaliana ribonucleoprotein mRNA, complete cds.
 ACCESSION M98340
 VERSION M98340.1 GI:166843
 KEYWORDS homologue; ribonucleoprotein; splicing factor; splicing factor SF-2.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1160)
 Lazar G.
 Unpublished (1992)
 Original source text: Arabidopsis thaliana (strain Landsberg erecta) (library: lambda gt11) root cDNA to mRNA.
 Location/Qualifiers
 1..1160
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Landsberg erecta"
 FEATURES
 source

```

BASE COUNT      300 a  233 c  318 g  309 t
ORIGIN

Alignment Scores:
Pred. No.:      2,34e-15      Length:      1160
Score:          139.00      Matches:      144
Percent Similarity: 66.67%      Conservative: 10
Best Local Similarity: 62.34%      Mismatches: 67
Query Match: 62.61%      Indels:    10
DB:              8      Gaps:      1

US-10-014-927-19MOD_COPY_1_222 (1-222) x ATHRIBNUCP (1-1160)

QY      1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
      |||
Db      31 ATGAGCAGTCGTCGAGTAGAACCGGTGACGTCGGAACTCTTCCTCGCGCATATCCGTGAG 90
QY      21 CysGluValGluAspLeupheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40
      |||
Db      91 AGAGAGGTCGAAGATTGTGTTCACTAGTATGGACCTGTTGTTCAAAATTGATTCAAGGTT 150
QY      41 ProProArgProProGlyTyrAlaPheValGluPheIuAspProArgAspAlaAsp 60
      |||
Db      151 CCTCAAGGCTCTCGTTATGATCATCTGTTGAGTTTGATGATCTCGGAGTGGCTGAAGAT 210
QY      61 AlalIeTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValcIuIleAla 80
      |||
Db      211 GCTATTCAATGCTCGTATGCTATGACTTTTGATGGCATCGTTTGAAGGTGGAAATTGGCG 270
QY      81 HisGlyGlyArgArg-----***** 90
      |||
Db      271 CATGTGGAGGCGGTTTCATCAGATGATCTCGGGGTAGTTTCATGGTGGTGGCGCGTGGT 330
QY      91 *****SerTyrSerAlaSerArgAlaProSerArgArgSerAspTyrArg 110
      |||
Db      331 GGTGTCGCGCGCGGTGGTGATGGTGCTAGTCTGGGCGCATCTAGGAGATCAGAGTTTCGT 390
QY      111 ValLeuValThrGlyLeuProProSerAlaSerTyrGluAspLeuLysAspHisMetArg 130
      |||
Db      391 GTTCTAGTCAATAGCTTGCTTTCATCTGCTCTTTGGCAAGATCTCAAGGATCACAATCGCT 450
QY      131 LysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyVal 150
      |||
Db      451 AAAGAGAGCGCATGCTGTTTCTCGAAGTGTACCGTGATGCTAGAGGGACAACCTGGAGTT 510
QY      151 ValAspTyrSerAsnTyrAspAspMetIysTyrAlaIleArgLysLeuAspAlaThrGlu 170
      |||
Db      511 GTTGATTTACCTCGCTATGAGGACATGAAGTATGCGCTGAAAAGCTCCACGACACAGAG 570
QY      171 PheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal 190
      |||
Db      571 TTTCCGAATTCGTTTTTCGAATGGATATGTCCTCGGGTTAGAGAAATATGATTCAGGAGGAGAT 630
QY      191 SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGlyProSer 210
      |||
Db      631 TCTAGAGGTCCTAGCCGGGGAAGATCCCTATTCTTAAGAGCCGACGCCGACCGCTGGACGA 690
QY      211 CysSerTyrSerSerLysSerArgSerValSer 221

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: February 4, 2004, 21:21:59 ; Search time 3510 Seconds
(without alignment)
2587.446 Million cell updates/sec

Title: US-10-014-927-19MOD_COPY_1_222
Perfect score: 222
Sequence: 1 MSRRNRRTIYVGNLFGDIRK.....RSRRGPGSCSYSSKSRVSVP 222

Scoring table: UNITARY2
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/BAUM927/runat_04022004_131333_2706/app_query.fasta_1.391
-DB=GenEmbl -QFWT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=unitary2 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=BAUM927 @CNG 1.1 3508 @runat_04022004_131333_2706 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_ov.*
5: gb_cv.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	222	100.0	838	8	AY150486 Arabidops
2	207	93.2	762	6	AX506504 Sequence
3	160	72.1	933	8	AY050912 Arabidops
4	143	64.4	858	8	BT006316 Arabidops
5	143	64.4	1229	8	AK118379 Arabidops
6	143	64.4	1232	8	AY085920 Arabidops
7	143	64.4	1540	8	AY128338 Arabidops
8	139	62.6	1160	8	M98340 Arabidops
9	131.5	59.2	1523	8	AY056185 Arabidops
10	131.5	59.2	2059	8	AK118074 Arabidops
11	95	42.8	1428	6	BD094070 Shear str
12	95	42.8	1428	9	M69040 Human SF2p3
13	95	42.8	2708	9	BC010264 Homo sapi
14	95	42.8	2765	9	AB062124 Homo sapi
15	95	42.8	2878	10	BC046773 Mus muscu
16	89	40.1	1618	5	BC042354 Xenopus l
17	88	39.6	1099	3	AK115991 Clona int
18	87.5	39.4	1069	6	AX410696 Sequence
19	87.5	39.4	1069	9	HSU30825 Human spli
20	87.5	39.4	1193	10	BC012217 Mus muscu
21	87.5	39.4	1584	6	AX305513 Sequence
22	87.5	39.4	1584	10	X66091 M.musculus
23	87.5	39.4	1717	9	M72709 Human alter
24	86.5	39.0	2369	5	BC046679 Xenopus l
25	73.5	33.1	1391	3	AF232773 Drosophil
26	73.5	33.1	1395	3	AY051919 Drosophil
27	68	30.6	809	3	AF242767 Caenorhab
28	68	30.6	160415	2	AC140961 Papio anu
29	65.5	29.5	343050	3	PFA929353 Plasmodiu
30	63	28.4	495	6	AX306308 Sequence
31	63	28.4	497	6	BD032870 Sequence
32	62	27.9	1485	9	BT007415 Homo sapi
33	62	27.9	1485	12	BT008254 Synthetic
34	62	27.9	2076	9	HUNSRP75A Human pre-m
35	62	27.9	2167	9	BC002781 Homo sapi
36	62	27.9	2204	10	BC019437 Mus muscu
37	59	26.6	1845	5	BC046895 Danio rer
38	58	26.1	3366	5	BC046668 Xenopus l
39	58	26.1	5164	8	AT131214 Arabidops
40	58	26.1	59261	8	T12M4 Arabidops
41	57.5	25.9	3081	5	BC044265 Xenopus l
42	57.5	25.9	3130	10	BC026944 Mus muscu
43	57	25.7	1579	10	RATERS Rattus norv
44	57	25.7	181545	10	AL626774 Mouse DNA
45	56	25.2	4044	6	AX040661 Sequence

ALIGNMENTS

RESULT 1

```

AY150486      838 bp      mRNA      linear      PLN 23-SEP-2002
LOCUS      Arabidopsis thaliana putative SP2/ASF splicing modulator Srp30
DEFINITION      (At1g09140) mRNA, complete cds.

ACCESSION      AY150486
VERSION      AY150486.1 GI:23297698
SOURCE      FLI_CDNA.
ORGANISM      Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 838)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J.,
Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T.,
Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Open Reading Frame (ORF) Clones
Unpublished
2 (bases 1 to 838)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J.,
Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T.,
Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.
FEATURES             source
     Location/Qualifiers
         1..838
            /organism="Arabidopsis thaliana"
            /mol_type="mRNA"
            /db_xref="taxon:3702"
            /chromosome="1"
            /clone="C10S028"
            /note="This clone is in pUNI 51.
            ecotype: Columbia"
         1..838
            /gene="At1g09140"
         1..807
            /gene="At1g09140"
         1..807
            /codon_start=1
         1..807
            /evidence=experimental
         1..807
            /product="putative SP2/ASF splicing modulator Srp30"
         1..807
            /protein_id="AA113011.1"
         1..807
            /db_xref="GI:23297698"
         1..807
            /translation="MSRRNRTIVGNLPDRLKCEVEDLFYKYGPIVDIDLKIPPRP
PGYAFVEFDRDADDAIYGRDGYDFDGLRVEIAHGGRRFSPVDYSSYSASRA
PSRRDYRLVITGLPPSASQDLKHKRAGDYCFSEVFPDRKMGVGVYDNDMK
YAIRKLDATFRNAPFSAYIRVREYSRVSQSPDDSKYSRGRSRGSPGSCYSKSR
SVSPARSISPRSLRSLRSLSVSRSSQSRKSRSRSRSSNSPVSPIVG"
         808..838
            /gene="At1g09140"
BASE COUNT      212 a      176 c      219 g      231 t
ORIGIN
1
3'UTR
1
Alignment Scores:
Pred. No.:      3,69e-32      Length:      838
Score:      222.00      Matches:      212
Percent Similarity:      100.00%      Conservative:      10
Best local Similarity:      95.50%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      8      Gaps:      0
US-10-014-927-19MOD_COPY_1_222 (1-222) x AY150486 (1-838)

QY      1 MetSerSerArgTTPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
DB      1 ATGAGTAGCCGATGGATCGTACGATCGTGGGAATTCGCTGGAGATATCGCAAG 60

21 CysGluValGluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuLysIle 40
DB      61 TGTGAGGTGAAGATCTCTCTTACAAGTATGACCAATTTGTGACATTTGATTGAAGATT 120

41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
DB      121 CCACCGAGACCTCTCTGTTATGCTTTGCGAGTTGAAGATCCCTCGTGATGCACACGAT 180

61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
DB      181 GCATTTATGGAGCTGATGTTATGATTTTGATGGGTTCGACTTCGGGTTCGAGATTGCA 240

81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100
DB      241 CATGGTGGTGGTGAAGATTTTTCACCATCAGTTGATAGGTACAGCAGCAGCTACAGTGCAGC 300

101 ArgAlaProSerArgArgSerArgTyrArgValLeuValThrGlyLeuProProSerAla 120
DB      301 CGTGCACCTTCAAGACGCTCTGACTACCGCGTGTGACCGGATTCACCGCTTCTCTCT 360

121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
DB      361 TCGTGGCAGGACCTTAAGGATCACATGCGCAAGCTGGAGATGTCGTCTTCTCTGAAGTT 420

141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
DB      421 TTCCTCGTACCGTGAAGCATGTCGGGTGTGGATTATAGCAACTATGATGATGATGAAG 480

161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
DB      481 TAGCAATAGGAACTTGATGCCACTGATTTCCAAATGCTTCTCTAGTCTTATATA 540

181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
DB      541 CGGGTCAGGGAATATGAGTCGAGGAGTGTGAGTGAAGCCAGATGATTTCTAAAGCTAT 600

201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
DB      601 AGAAGCAGGAGTCGGAGCCGTGTCGACTGTAGCTATAGTAGCAGGAGGAGTGTG 660

221 SerPro 222
DB      661 TCACCT 666

RESULT 2
AX506504      762 bp      DNA      linear      PAT 27-SEP-2002
LOCUS      Sequence 1199 from Patent WO216655.
DEFINITION      AX506504
ACCESSION      AX506504
VERSION      AX506504.1 GI:23387741
KEYWORDS      Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1
Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
Patent: WO 0216655-A 1199 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG
(CH)
Location/Qualifiers
1..762
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT      189 a      161 c      197 g      215 t
ORIGIN
1
Alignment Scores:
Pred. No.:      2.86e-29      Length:      762

```

Score: 207.00 Matches: 197
Percent Similarity: 100.00% Conservative: 10
Best Local Similarity: 95.17% Mismatches: 0
Query Match: 93.24% Indels: 0
DB: 6 Gaps: 0

US-10-014-927-19MOD_COPY_1_222 (1-222) x AX506504 (1-762)

QY 1 MetSerArgTrpAsnArgThrIleTyrValGlyValAsnLeuProGlyAspIleArgLys 20
Db 1 ATGAGTAGCCGATGGAATCGATCGATCTACGATCTACGTTGGAAATTCCTGGAGATATTCGCAAG 60
QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProlIleValAspIleAspLeuLysIle 40
Db 61 TGTGAGGTTGAAGATCTCTTACAAAGTATGGACCAATTTGGACATTTGATTGAAGATT 120
QY 41 ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
Db 121 CCACCGAGACCTCTGTTGATGCTTTGCGAGTTTGAAGATCCTCGTATGATCGACGAT 180
QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db 181 GCAATTTATGACGATGATGTTTATGATTTGATGGTGTGCGACTTCGGGTTGAGATTGCA 240
QY 81 HisGlyGlyArgArg*****SerTyrSerAlaSer 100
Db 241 CATGGTGTGTAGATTTTCCACCATCAGTTGATAGGTACAGCAGCAGCTACAGTTCGACG 300
QY 101 ArgAlaProSerArgSerArgTyrArgValLeuValThrClyLeuProProSerAla 120
Db 301 CGTGACCTTCAAGACGCTCTGACTACCGGCTGCTTTGACCGGATTCGCGCTTCTGCT 360
QY 121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
Db 361 TCGTGGCAGACCTTAAAGATCATCGCAAGCTGGAGATGCTCTGCTTCTGAGTT 420
QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
Db 421 TTCCCTGACGTAAGACATGCTGGGTTGTGGATATAGCAACTATGATGATATGAAG 480
QY 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
Db 481 TAGCAATAAGAACTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
Db 541 CGGCTGAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 201 ArgSerArgSerArgSerArg 207
Db 601 AGAAGCAGGAGTCGGAGCCGT 621

RESULT 3
AY050912 933 bp mRNA linear PLN 18-SEP-2002
LOCUS Arabidopsis thaliana putative SP2/ASF splicing modulator Srp30
DEFINITION protein (At1g09140) mRNA, partial cds.

ACCESSION AY050912
VERSION AY050912.1 GI:15292956
KEYWORDS FLI CDNA
SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 933)
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banno, J., Chung, M.K.,
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,
Davis, R.W., Ecker, J.R. and Theologis, A.

Arabidopsis Full Length cDNA Clones

Unpublished
2 (bases 1 to 933)
Yamada, K., Liu, S.X., Pham, P.K., Banno, J., Banno, F., Dale, J.M.,
Goldsmith, A.D., Jiang, P.X., Lee, J.M., Chodera, C.S., Quach, H.L.,
Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C.,
Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M.,
Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P.,
Southwick, A., Tracy, S.E., Shinozaki, K., Davis, R.W., Ecker, J.R. and
Theologis, A.
Direct Submission
Submitted (24-JUL-2001) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PDEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Yamada, K., Liu, S.X.,
Pham, P.K., Banno, J., Banno, F., Dale, J.M., Goldsmith, A.D.,
Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,
Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H.,
Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B.,
Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P.,
Southwick, A., Tracy, S.E., Davis, R.W., Ecker, J.R. and Theologis, A.
Yamada, K. (SSP/PDEC) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PDEC)
contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.

Location/Qualifiers
1..933
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL08-12-105 (R11038)"
/note="This clone is in a modified pBluescript vector
(FLC-1) as a BamHI/XhoI insert."
ecotype: Columbia
c1..933
/gene="At1g09140"
c1..626
/gene="At1g09140"
/codon_start=3
/evidence=experimental
/product="putative SP2/ASF splicing modulator Srp30
protein"
protein_id="AAK33589.2"
/db_xref="GI:21954085"
/translation="TYRSDGYDFDQCLRLVEIAHGRRFSPVDRYSSYSASRAPSR
RSDHYLVLTGLPPSASMDLADHMKAGDVCSEVFPDRKMGSDVYDYNDDMKYAI
KLDKATEFRNFASSAIYRVEYSESRSPDDSKSYSESRSRGSPGSCSYSKSRVS
PARSISPSRFLSRSLSYSSVRSQSKSRKTRSRNSPSPVISG"

misc_difference 1
/gene="At1g09140"
/note="compared to genomic sequence"
/replace="c"
misc_difference 148
/gene="At1g09140"
/note="compared to genomic sequence"
/replace="g"
misc_difference 579
/gene="At1g09140"
/note="compared to genomic sequence resulting in an amino
acid sequence difference"
/replace="t"

Db	61	AGAGAGGTCGAGGATTTTCTTCAGTAAGTATGAGCACTCTTCTTCAAAATTTGATTGAGGTTT	120
Qy	41	P-oPrAcArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp	60
Db	121	CCTCCAAAGCCTCTCCGGTATGCAATTTGTTGAGTTTCATGATCTCGGATCTGAGAT	180
Qy	61	AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla	80
Db	181	GCTATTTCATGTCGTGATGGCTATGACTTTTTCATGGGCATGCTTTGAGGGTGAATTTGGCG	240
Qy	81	HisGlyGlyArgArg-----*****	90
Db	241	CATGTGGAGCGCTTCATCAGATCATCTCGGGTAGTTTCAATGGTGGCCGTGGT	300
Qy	91	*****SerTyrSerAlaSerArgAlaProSerArgArgSerAspTyrArg	110
Db	301	GTTGTCGGCGCCGGTGTATGGTGTATGTCGTGGCCATCTAGAGATCAGATTCGT	360
Qy	111	ValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuIysAspHisMetArg	130
Db	361	GTTCTAGTCACAGGCTTCCCTTCATCTCTTCTTGGCAAGATCTCAAGGATCACA	420
Qy	131	LysAlaGlyAspValCysPheSerGluValPheProAspArgIysGlyMetSerGlyVal	150
Db	421	AAAGAGAGCGCATGTCGTCTTTTCGCAAGTGACCGTGATCTAGAGGGCAACTGGAGTT	480
Qy	151	ValAspTyrSerAsnTyrAspAspMetIysTyrAlaIleArgLysLeuAspAlaThrGlu	170
Db	481	GTTGATTACCTGCTATGAGCATGAGTATGCGCTGAAAAGTCTGACGACACAGAG	540
Qy	171	PheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal	190
Db	541	TTTTCGAATTCGTTTTTCGAATGGATATGTCGGGTAGAGATATGATTCACAGGAAGGAT	600
Qy	191	SerArgSerProAspAspSerIysSerTyrArgSerArgSerArgSerArgGlyProSer	210
Db	601	TCTAGAGTCTTAGCGGGGAGAGTCCCTATTCTAAGAGCCGACGCCGACCGTGGACGA	660
Qy	211	CysSerTyrSerSerIysSerArgSerValSer	221
Db	661	ACGCTGAGCCGAGCAGGAGCAGAGAGCAGGAGC	693
RESULT 5			
AK118379			
LOCUS			
DEFINITION	AK118379	1229 bp mRNA linear PLN 06-DEC-2000	
ACCESSION	AK118379.1	GI:26451793	
VERSION	AK118379.1	GI:26451793	
KEYWORDS	FLI CDNA; CAP trapper.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et al. (2001) Genomics 77:79-90) digested with BamHI and SalI. This clone is in a modified pBluescript vector.

Please visit our web site (<http://pfjweb.gsc.riken.go.jp/>) for further details.

FEATURES

source Location/Qualifiers

1. .1229

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="1"

/clone="RAF119-64-H16"

/notes="common name: thale cress"

1. .1229

/gene="At1G02840/F22D16_30"

93. .1004

/gene="At1G02840/F22D16_30"

/codon_start=1

/product="putative ribonucleoprotein SF-2"

/protein_id="BAC42991.1"

/db_xref="GI:26451794"

/translation="MSSRSRSTVYVGNLPGDIRVEDLFSKYGPVQVLDLKVPPRP
PGVAFVFDADAEDAIHGRDGYDFDHRLRVELAHGRRSSDDTGTSGFNGSGRG
RGGDGGSPRRSRFRVLVLDLPSSASWDLKDHWRKGDVCFQGVYDRADTGV
VDYTCEDMKYALKLDDTFRNFNGYVRVREYDKDSDSPSRGSRYSKRSR
GRSVRSRSPSRSPRAKSRSPAKSTSPGPRSKSRSPSRSPSRSPSRSP
SVQKEGSKSPSPSPAKSPPIHNSPSR"

BASE COUNT 318 a 256 c 328 g 327 t

ORIGIN

Alignment Scores:

Pred. No.: 4,26e-16 Length: 1229

Score: 143.00 Matches: 148

Percent Similarity: 68.40% Conservative: 10

Best Local Similarity: 64.07% Mismatches: 63

Query Match: 64.41% Indels: 10

DB: 8 Gaps: 1

US-10-014-927-19MOD_COPY_1_222 (1-222) x AK118379 (1-1229)

Qy 1 MetSerSerArgTrpIenArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20

Db 93 ATGAGCAGTCGTTCCGAGTAGAACCGTGATCGTGGAAACCTTCCTGCGCATATCCGTGAG 152

Qy 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40

Db 153 AGAGAGATCGAAGATTGTTTCAGTAAGTATGGACCTGTTGTTCAAAATTGATTGAGGTT 212

Qy 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60

Db 213 CCTCCAGGCGCTCTGTTATGATTGATTCGTTGAGTTGATGCTCGGATGCTGAAGAT 272

Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80

Db 273 GCTATTATGTCGTGATGCTATGACTTGTATGGCATCGTTTGAAGGTCGAAATGCGC 332

Qy 81 HisGlyClyValArg-----***** 90

333 CATGTCGGAGGCGTTTCATCAGATGATCTCGGGGTAGTTTCAATGTTGGTGGCCGTGGT 392

Qy 91 *****SerTyrSerAlaSerArgAlaProSerArgArgSerAspTyrArg 110

Db 393 GGTGTCGGCGCGGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 452

Qy 111 ValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArg 130

Db 453 GTTCTAGTCACAGGCTTGCCTTCATCTGCTTCTTGGCAAGATCTTCAGAGGATCATCGCT 512

Qy 131 LysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyVal 150

Db 513 AAAGAGGCGCATGTCCTGTTCTCCCAAGTGTACCGTATGCTAGAGGGAACAATGGAGTT 572

QY	151	ValAspTyrSerAsnTyrAspMetLysTyrAlaIleArgLysLeuAspAlaThrGlu	170
Db	573	GTGTGATTACACTGCTATGAGACATGAATGATCGCTGAAAGAGCTCGACGACAGAG	632
QY	171	PheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal	190
Db	633	TTTCGAAATGGCTTTTCGAATGGATATGTCGGGTGTAGAGATATGATTCAAGGAAGAT	692
QY	191	SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGlyProSer	210
Db	693	TCTAGAGTCTAGCGCGGAAGATCTTATTCTAAGAGCCGACGCGCGGAGACGA	752
QY	211	CysSerTyrSerSerLysSerArgSerValSer	221
Db	753	AGCGTGAGCGAAGCAGGACGAGACGAGCAGC	785
RESULT 6			
AY085920			
LOCUS			
Arabidopsis thaliana clone 19681 mRNA, complete sequence.			
ACCESSION			
AY085920.1			
VERSION			
AY085920.1			
KEYWORDS			
FLI CDNA.			
SOURCE			
Arabidopsis thaliana (thale cress)			
ORGANISM			
Arabidopsis thaliana			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
REFERENCE			
1 (bases 1 to 1232)			
Haas,B.J., Volfovsky,N., Town,C.D., Troughan,M., Alexandrov,N.,			
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.			
Full-length messenger RNA sequences greatly improve genome			
annotation			
Genome Biol. 3 (6), RESEARCH0029 (2002)			
JOURNAL			
MEDLINE			
PUBMED			
22088475			
12093376			
REFERENCE			
2 (bases 1 to 1232)			
Brover,V., Troughan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and			
Feldmann,K.			
Full-length cDNA from Arabidopsis thaliana			
Unpublished			
REFERENCE			
3 (bases 1 to 1232)			
Brover,V., Troughan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and			
Feldmann,K.			
Direct Submission			
Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,			
Malibu, CA 90265, USA			
COMMENT			
This clone sequence is one of 5,000 Ceres full-length cDNAs made			
available to TIGR and Genbank. The following quality assessment of			
this set was done by comparison with known proteins: two percent of			
the clones are estimated to be 5'-truncated; less than one percent			
are 3'-truncated; approximately two percent represent alternative			
splice variants, including unspliced introns and spliced exons; one			
percent may contain premature stop codons; five percent may have			
frame shifts in a coding region. A sequence is considered to be			
5'-truncated if it lacks the translation initiation start (ATG). A			
sequence is considered to be 3'-truncated if it lacks the			
C-terminal end of the encoded protein. Please note that these cDNA			
sequences are derived from the ws or Laer ecotypes and therefore			
may contain polymorphisms when compared to sequences from Col-0.			
Genset carried out the library production and sequencing of the			
full-length clones. Ceres, Inc. carried out the clustering of the			
5' sequences, selection of clones, and sequence assembly.			
Location/Qualifiers			
1. .1232			
/organism="Arabidopsis thaliana"			
/mol_type="mRNA"			
/db_xref="taxon:3702"			
/clone="19681"			
94..1005			
/codon_start=1			
/product="SF2/ASF-like splicing modulator Srp30, putative"			
CDS			
FEATURES			
source			
1. .1232			
/organism="Arabidopsis thaliana"			
/mol_type="mRNA"			
/db_xref="taxon:3702"			
/clone="19681"			
94..1005			
/codon_start=1			
/product="SF2/ASF-like splicing modulator Srp30, putative"			
CDS			
RESULT 7			
AY128338			
LOCUS			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			
AY128338			
1540 bp			
mRNA			
linear			
PLN 07-AUG-2002			
Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,			
DEFINITION			

putative (At1g02840) mRNA, complete cds.

AY128338
VERSION AY128338.1 GI:22135917
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1540)
Tripp,M., Southwick,A., Karlin-Neumann,G., Nguyen,M., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,Y., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (01-JUL-2002) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arab@sequence.stanford.edu

The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL CDNA: 'RIKEN
Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and
sequenced the pUNI (ORF) clones using the RAFL cDNAs: Tripp,M.,
Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R.,
Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H.,
Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Tripp,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally
to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.

FEATURES
source
1..1540
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL04-19-E10 (R15578)"
/note="This clone is in pBluescript
ecotype: Columbia"
1..1540
/gene="At1g02840"
/note="synonym: F22D16.16"
49..906
/gene="At1g02840"
/codon_start=1
/product="SF2/ASP-like splicing modulator Srp30, putative"
/protein_id="AA091541.1"
/db_xref="GI:22135918"
/translation="MSRSRTYVGNLPDIEREVEDLFSKYGPVQIDLVKPPRP
PGVAFEDDARDAEDAIHGDYDFGHLRVELAHGRRSSDDTRGSFNGGRRGG
RGDGSGSGSRSEPRVLVTGLPSSAWQDLKRMKGGDYCFQVYRDAGTTGV
VDYTCYDMKYLKLLDTEFRNFGVYVREYDSDRSRSPRSRGRYSRSPSR
GRSVRSRSPSRSPSPKASRSRSPKSTSRSPGSRSPSRSPSRSPSRSPSRSPSR
DHNIIISGL"

BASE COUNT 398 a 399 g 446 t
ORIGIN

Alignment Scores:
Pred. No.: 7,01e-16 Length: 1540
Score: 143.00 Matches: 148
Percent Similarity: 68.40% Conservative: 10
Best Local Similarity: 64.07% Mismatches: 63
Query Match: 64.41% Indels: 10
DB: Gaps: 1

US-10-014-927-19MOD_COPY_1_222 (1-222) x AY128338 (1-1540)

QY 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
DB 49 ATGAGCAGTCGTTGAGTAGAACCGTGTACGTCGGAACCTTCCTGCGGATATCCGTGAG 108
QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40
DB 109 AGAGAGGTCGAAGATTGTTTCAGTAATGATGACCTGTTGTTCAAATTGATTGAAGTT 168
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60
DB 169 CCTCCAAAGCCCTCCTGGTTATGCAATTCGTTGAGTTGATGCTCGGGATGCTGAAGAT 228
QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
DB 229 GCTATTTCATGTCGTGATGCTGATGCTATGACTTTCATGCGCATCGTTTCAGGGGTGGAATTGCGC 288
QY 81 HisGlyGlyArgArg-----***** 90
DB 289 CATGTTGGGAGCGGTCATCAGATGACTCGGGGTAGTTTCAATGTTGGTGGCGGTGGT 348
QY 91 *****SerTyrSerAlaSerArgAlaProSerArgArgSerAspTyrArg 110
DB 349 GGTGTCGCGCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 408
QY 111 ValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArg 130
DB 409 GTTCTAGTCACAGGCTTGCCTTCATCTGCTTCTTGGCAAGATCTCAAGGATCACATGCGT 468
QY 131 LysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyVal 150
DB 469 AAAGAGAGCGATGTCTGTCTCTCGCAAGTGTACCGTGTAGTAGGAGCAACTGGAGTT 528
QY 151 ValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGlu 170
DB 529 GTTGATTACCTGCTATGAGGACATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 588
QY 171 PheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal 190
DB 589 TTTCCAAATGGGTTTCCGAATGATATGTCGGGTAGAGAAATATGATTCAAGGAAGAT 648
QY 191 SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGlyProSer 210
DB 649 TCTAGGAGTCCTAGCGCGGGAAGATCTTATTCTAAGACCGCAGCGCGCGCGGTGGACGA 708
QY 211 CysSerTyrSerSerLysSerArgSerValSer 221
DB 709 ACCGTGACCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 741

RESULT 8
ATHRIENUCP 1160 bp mRNA linear PLN 27-APR-1993
LOCUS Arabidopsis thaliana ribonucleoprotein mRNA, complete cds.
DEFINITION M98340
ACCESSION M98340
VERSION M98340.1 GI:166843
KEYWORDS homologue; ribonucleoprotein; splicing factor; splicing factor
SF-2.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1160)
AUTHORS Lazar,G.
JOURNAL Unpublished (1992)
COMMENT Original source text: Arabidopsis thaliana (strain Landsberg erecta) (library: Lambda gt11) root cDNA to mRNA.
FEATURES
source
1..1160
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Landsberg erecta"

1026)

Db 213 AGAAGAGTTGAAGACTTGTTCAGTAAAGTATGACCTGTGTTCAAATCGATTGAAGATT 272
QY 41 PProAATGProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
Db 273 CGCGGAGGCTCCAGGCTATGATCGTCGAGTTTGAAGATGCTGCGATGCTGATGAT 332
QY 61 AlalleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db 333 GCAATTATGCGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 392
QY 81 HisGlyGlyArg-----***** 89
Db 393 CATGTTGGGAGCGTTTCATCATCATGATGACCGGCTAGTATTAGTGTGGTGGTGGC 452
QY 90 *****SerTyrSerAlaSerArgAlaPProSerArgSerArgSerArgTyr 109
Db 453 GTCCTGT 512
QY 110 ArgValLeuValThrGlyLeuProProSerAlaSerTyrGlnAspLeuLysAspHisMet 129
Db 513 CGCGTTGTAGTTCAGGTTTCTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 572
QY 130 ArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGly 149
Db 573 CGTAAGGAGGAGAAAGTTTCTCAAGTGTTCGTGATGATGATGATGATGATGATGATG 632
QY 150 ValValAspTyrSerAsnTyrAspMetLysTyrAlaIle-ArgLysLeuAspAlaLth 169
Db 633 ATTGTAGATTATACAGCTACAGAGCATGAATATGCGATAAAGAGCTCGATGACAC 692
QY 169 rGluPheArgAsnAlaPheSerSerAlaTyrIleArgValrGluTyrGluSerArgse 189
Db 693 AGAGTTTCGGAATGCGTTTCTCATGATATGTTCTCGGTTAGAGATATGATGATGATG 752
QY 189 rValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGlyPr 209
Db 753 GGATTCGAGAGCGCGCGCGGAGGATCTTCTTAAGAGCGCGCGCGCGCGCGCGCGGAG 812
QY 209 oSerCysSerTyrSerLysSerArgSerValSer 221
Db 813 TCCAGCGCTAGTCTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 849

BD094070 1428 bp DNA linear PAT 27-AUG-2002
LOCUS Shear stress-responsive DNAs.
DEFINITION BD094070
ACCESSION BD094070.1 GI:22639658
VERSION WO 0125427-A/31.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1428)
AUTHORS Nijima,H., Yoshisue,H., Obayashi,M., Ota,T., Kawabata,A.,
Sakurada,K., Kuga,T., Sekine,S., Nakamura,Y. and Sugan,S.
TITLE Shear stress-responsive DNAs
JOURNAL Patent: WO 0125427-A 31 12-APR-2001;
KYOWA HAKKO KOGYO CO LTD,HIROSHI NOJIMA,HAJIME YOSHISUE, MASAYA
OBAYASHI, TOSHIO OTA, AVAKO KAWABATA,KAZUHIRO SAKURADA, TETSURO KUGA,
SUSUMU SEKINE, YUSUKE NAKAMURA, SUMIO SUGANO
COMMENT OS Homo sapiens (human)
PN WO 0125427-A/31
PD 12-APR-2001
PF 02-OCT-2000 WO 2000JP006840
PR 01-OCT-1999 JP 99P 280976
PI HIROSHI NOJIMA,HAJIME YOSHISUE,MASAYA OBAYASHI,TOSHIO OTA, PI
AYAKO KAWABATA,
PI KAZUHIRO SAKURADA,TETSURO KUGA,SUSUMU SEKINE,YUSUKE NAKAMURA,
SUMIO SUGANO
PC C12N15/12,C07K14/435,C07K16/18,C12P21/02,C12Q1/68,A61K38/00,
A61K39/395,
PC

PC A61K48/00,A61P9/10,G01N33/50,G01N33/53
CC
FH Key Location/Qualifiers
FT CDS (125)..(868).

FEATURES
source
1..1428
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 341 a 296 c 400 g 391 t
ORIGIN

Alignment Scores:
Pred. No.: 2,04e-06 Length: 1428
Score: 95.00 Matches: 108
Percent Similarity: 56.46% Conservative: 10
Best Local Similarity: 51.67% Mismatches: 85
Query Match: 42.79% Indels: 6
DB: Gaps: 2

US-10-014-927-19MOD_COPY_1_222 (1-222) x BD094070 (1-1428)

QY 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
Db 176 ATCTAGTGGTAACTTACCTCCAGACATCCGACATCCGAAACCAAGACATTGAGACGTGTTCTAC 235
QY 29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProGlyTyrAla 48
Db 236 AANTACGGCGCTATCCGACATCCGACATCCGACATCCGACATCCGACATCCGACATCCGAC 295
QY 49 PheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyr 68
Db 296 TTCGTTGAGTTCGAGACCCGCGAGACGCGGAGACGCGGTGTATGTCGCGACGCGCTAT 355
QY 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArg-----Arg 85
Db 356 GATTACGATGGGTACCGTCTCGCGGTGGAGTTTCCTCGAAGCGCGCGCTCGAAGACGCCGA 415
QY 86 *****SerTyrSerAlaSerArgAlaProSerArg 105
Db 416 GCGCGCGCGGGGTGGAGTGGCGGAGTCCCGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 475
QY 106 ArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeu 125
Db 476 CGGTCTGAAAACAGAGTGTGTCTCTGACATCCCTCAAGTGGAAAGTTGGCAGGATTTA 535
QY 126 LysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLys 145
Db 536 AAGGATCACATCGTGAAGCAGGTGATGTATGTATGTATGTATGTATGTATGTATGTATGTAT 589
QY 146 GlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLys 165
Db 590 ---GGCAGCTGGTGTCTGGAGTTGTACGGAAAGAGATATGACCTATGCGATTGCAAAA 646
QY 166 LeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyr 185
Db 647 CTGGATAACACTAAGTTTAGATCTCATGAGGGGAAACTGCCTACATCGGGTTAAAGTT 706
QY 186 GluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArg 205
Db 707 GATGGGCCCAAGAGTCCAAAGTTATGAGATCTCGATCTCGAAGCGGTAGTCTGATGATGAT 766

QY 206 SerArgGlyProSerCysSerTyrSer 214
Db 767 AGCCGTAGCAGAGCAACAGCAGGAGT 793

RESULT 12
HUMSF2P33 1428 bp mRNA linear PRI 24-NOV-1993
LOCUS HUMSF2P33
DEFINITION Human SF2p33 mRNA, complete cds.
ACCESSION M69040
VERSION M69040.1 GI:338046
KEYWORDS SF2p33.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Krainer, A.R., Mayeda, A., Kozak, D., and Binns, G.
TITLE Functional expression of cloned human splicing factor SF2: homology
to RNA-binding proteins, UI 70K, and drosophila splicing regulators
Cell 66 (2), 383-394 (1991)
JOURNAL 91309150
MEDLINE
PUBMED 1830244
COMMENT Original source text: Homo sapiens cDNA to mRNA.
Comments: SF2p33 is an essential pre-mRNA splicing factor, which
can also affect alternative 5' splice site selection in vitro by
stimulating the use of proximal 5' splice sites.
(Krainer, A.R., Conway, G.C., and Kozak, D. (1990) Purification and
Characterization of SF2, a Human Pre-mRNA Splicing Factor. Genes
Dev. 4, 1158-1171; Krainer, A.R., Conway, G.C., and Kozak, D.
(1990) The Essential Pre-mRNA Splicing Factor SF2 Influences 5'
Splice Site Selection by Activating Proximal Sites. Cell 62,
35-42). This factor is also known as ASF (Ge, H., and
Manley, J.L. (1990) A protein factor, ASF, Controls Alternative
Splicing of SV40 Early Pre-mRNA in Vitro. Cell 62, 25-34). SF2 p33
consists of two polypeptides of apparent molecular weight 33 kd,
approximately, although their predicted molecular weight is 27,744
daltons. The two forms appear to differ by the extent of
post-translational modification, which includes phosphorylation.
SF2p33 binds RNA and promotes the annealing of complementary RNAs.
It is required for assembly of pre-spliceosome complexes. The
N-termini of the HeLa polypeptides are blocked.
Location/Qualifiers
1. 1428
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="female"
125..871
/functions="pre-mRNA splicing factor"
/codon_start=1
/product="SF2p33"
/protein_id="AAA03476.1"
/db_xref="GI:338047"
TRANSLATION="MSGGGVIRPAGNDCRIYVGNLPDIPDIEDVYKYGAIRD
IDLKNGGGPPFAPVFEDPRDAEDAVYGDYDYGVRILRVFPSSRGRTGGGGG
GGGAGRGVGPSPRSRSENVVYSGLPFGSGWQDLKHREAGDVYADYRDGVV
EFVKEDMTAVAKLQNTKPRSEGETAYIRVKGDPSPRSRSGRSRSGRSR
SNRSRSYSPSRSGSPRYSPRSRST"
872..1428
3'UTR
BASE COUNT 341 a 296 c 400 g 391 t
ORIGIN
Alignment Scores:
Pred. No.: 2,04e-06 Length: 1428
Score: 95.00 Matches: 108
Percent Similarity: 56.46% Conservative: 10
Best Local Similarity: 51.67% Mismatches: 85
Query Match: 42.79% Indels: 6
Gaps: 2
US-10-014-927-19MOD_COPY_1_222 (1-222) x HUMSF2p33 (1-1428)
Qy 9 IleTyrValGlyAenLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
Db 176 ATCTACGTGGGTACTTACCTCCAGACATCGCAACCAAGACATTGAGGACGTGTTCTAC 235
Qy 29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla 48
Db 236 AAATACGGCGCTATCCCGGACATCGACTCAAGATCGCGCGGGGACGCCCTTCGCC 295
Qy 49 PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr 68
Db 296 TTCGTTAGTTCGAGGACCGCGAGACCGCGGAGACCGCGGTGTATGTCGCGACGCTAT 355

Qy 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArg-----Arg 85
Db 356 GATTACGATGGTACCGGTCCTCGGGTGGAGTTTCTCGAAGCGCCGCTGGGAACAGGCGGA 415
Qy 86 *****SerTyrSerAlaSerArgAlaProSerArg 105
Db 416 GCGGCGCGGGGTGGAGTGGGGAGCTCCCGAGGTGCTATGGCCCCCATCCAGG 475
Qy 106 ArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTyrGluAspLeu 125
Db 476 CGGTCTGAAACACAGAGTGGTTGTTCTCTGACTGCTCCAGTGGGAAGTTGCGAGATT 535
Qy 126 LysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLys 145
Db 536 AAGGATCATCATCGGTGAGCAGGTGATGTTATGCTGATGTTTACCGAGAT----- 589
Qy 146 GlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLys 165
Db 590 ---GGCATCTGGTGTGGAGTTTGTACGGAAGAGATATGACCTATGCACTTCGAA 646
Qy 166 LeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyr 185
Db 647 CTGGATACATAGTTTAGATCTCATGAGGAGAACTGCTACATCCGGGTTAAAGTT 706
Qy 186 GluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArg 205
Db 707 GATGGGCCAGAGTCCCAAGTTATGGAAGATCTCGATCTCGAAGCCGTAGTCGTAGCAGA 766
Qy 206 SerArgGlyProSerCysSerTyrSer 214
Db 767 AGCCGTAGCAGAAGCAACAGCAGGAGT 793

RESULT 13
LOCUS BC010264
DEFINITION Homo sapiens, splicing factor, arginine/serine-rich 1 (splicing factor 2, alternate splicing factor), clone MGC:5228 IMAGE:2900101, mRNA, complete cds.
ACCESSION BC010264
VERSION BC010264.1 GI:16307433
SOURCE MGC.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2708)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalon@bcm.tmc.edu
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IMAGE Plate: 3 Row: e Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5902075.

102(b) 10

FEATURES	Location/Qualifiers	
	1..2708	
source	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="LocusID:6426"	
	/db_xref="taxon:9606"	
	/clone="MGC:5228 IMAGE:2900101"	
	/tissue_type="Placenta, choriocarcinoma"	
	/clone_lib="NIH MGC_10"	
	/lab_host="DH10B"	
	/note="Vector: pCMV-SPORT6"	
	36..782	
CDS	/codon_start=1	
	/product="splicing factor, arginine/serine-rich 1 (splicing factor 2, alternate splicing factor)"	
	/protein_id="AAH10264.1"	
	/db_xref="GI:16307434"	
	/translation="MSGGVIRGPGAGNDCRIYVGNLPPDITKDIEDVFYKYGAIRD IDLKNRGGPPAFVFEEDPRDAEDAVYDGRDLYDGRVLRVFEPRSGRGTGGGGG GGGAPRGYGPSSRNVRVVGSPGQDLKDHREAGDVCVADVVDGTGVV EFVRKEDMTYAVRKLDNTKFRSHEGETAIRVKVDGPRSPSYGRSRSR SNRSRSYSPRSRSGSPYSPRHSRSRT"	
	754 a	485 c 627 g 842 t
BASE COUNT		
ORIGIN		
Alignment Scores:		
Pred. No.:		
Score:		
Percent Similarity:		
Best Local Similarity:		
Query Match:		
DB:		
US-10-014-927-19MOD_COPY_1_222 (1-222) x BC010264 (1-2708)		
QY	9	IleTyrValGlyAsnLeuProGlyAspLeuArgLysCysGluValGluAspLeuPheTyr 28
Db	87	ATCTAGTGGTAACTTACTCCAGACATCCGAACCAAGGACATTCGAGCGTGTCTTAC 146
QY	29	LysTyrGlyProIleValAspLeuAspLeuLysIleProProArgProGlyTyrAla 48
Db	147	AAATACGGCGCTATCCGCGACATCCACCTCAGAGATCCGCGGGGACCGCCCTTCGCC 206
QY	49	PheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyr 68
Db	207	TTCGTGAGTTCCAGACCCCGGAGACCGCGGAGCGGGTGTATGTCGCGAGCGCTAT 266
QY	69	AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArg-----Arg 85
Db	267	GATTACGATGGGTACCGTCTCGGGTGGAGTTTCCTCGAAGCGCGCGTGAACAGGCCGA 326
QY	86	*****SerTyrSerAlaSerArgAlaProSerArg 105
Db	327	GGCGGCGGGGGTGGAGTGGCGGAGCTCCCGAGGTGCGTATGCGCCCGCCATCCAGG 386
QY	106	ArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTyrGluAspLeu 125
Db	387	CGGTCTGAAACACAGAGTGTGTCTCGACTGCCTCCAAAGTGAAGTTCGAGATTTA 446
QY	126	LysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLys 145
Db	447	AAGATCAATGCGTGAACAGGTGATGTATGTTATGCTGATGTTTACCGAGAT----- 500
QY	146	GlyMetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAlaIleArgLys 165
Db	501	---GGCACTGGTCTCGTGGAGTTTGTACGGAAGAAGATATGACCTATGCAGTTCGAAA 557
QY	166	LeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyr 185
Db	558	CTGGATAACATAAGTTTATAGTCTCATGAGGAGAAATCGCTATACCGGGTTAAAGTT 617
QY	186	GluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArg 205

QY 69 AspPheAspGlyCysArgLeuArgValGluLeuAlaHisGlyGlyArg-----Arg 85
Db 322 GATTACAGTGGTACCGTCTCGGGTGGAGTTCTCGAAGCGCGCCGTGGACAGCGCGA 381
QY 86 *****SerTyrSerAlaSerArgAlaProSerArg 105
Db 382 GCGCGCGCGGGGTGAGGTGGCGAGTCCCGAGGTGCTATGCGCCCGCCATCCAGG 441
QY 106 ArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeu 125
Db 442 CGGTCTGAACACAGAGTGGTGTCTCGGAGTGGCTCCCAAGTGAAGTGGCAGGATTTA 501
QY 126 LysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLys 145
Db 502 AAGGATCACATGCTGAGCAGGTGATGTTATGCTGATGTTATCCAGAT----- 555
QY 146 GlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaLeuArgLys 165
Db 556 ---GGCACTGGTGGCGAGTGGTGGTGGCGAAGAGATATGACCTATGAGTTCGAA 612
QY 166 LeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyr 185
Db 613 CTGATACACTAAGTTAGTCTCATGAGGAGAACTGCTACATCCGGTTAAAGTT 672
QY 186 GluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArg 205
Db 673 GATGGCCCAAGTCCCAAGTTATGGAAGATCTCGATCTCGAAGCGGTAGTGTAGCAGA 732
QY 206 SerArgGlyProSerCysSerTyrSer 214
Db 733 AGCCGTAGCAGACAGCAACAGCAGGAGT 759

RESULT 15

BC046773 2878 bp mRNA linear ROD 14-FEB-2003
LOCUS
DEFINITION Mus musculus, Similar to splicing factor, arginine/serine-rich 1 (splicing factor 2, alternate splicing factor), clone MGC:61417 IMAGE:5708288, mRNA, complete cds.

ACCESSION

VERSION BC046773.1 GI:28386235

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2878)
Strausberg, R.
Direct Submission
Submitted (13-FEB-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REFERENCE

AUTHORS
TITLE
JOURNAL
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldi, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Finkler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

REMARK

COMMENT
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: Plate: Row: Column: 0
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF

analysis, Genomescan gene prediction.

FEATURES

source

1..2878
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:61417 IMAGE:5708288"
/tissue_type="Brain, mouse 15.5 dpc"
/clone_lib="NIH BMAP_EW0"
/lab_host="DH10B"
/note="Vector: pYX"
99..845
/codon_start=1
/product="Similar to splicing factor, arginine/serine-rich 1 (splicing factor 2, alternate splicing factor)"
/protein_id="AAH46773.1"
/db_xref="GI:28386235"
/translation="MSGGVVRGAGNNDRIYVGNLPDITKIDIEVFKYGAIRD IDLKNRRGGPFVAFVEFDPDASDAVYDRGDYDYGRLRVEFPRSGRGTRGGGGV GGGAGPRGYGPPRRSRNRYVSGLPSPSGQDLKDMRAGDVYADVYRDGGVY EFVRKEDMTYAVRKLDTKFRSHGSETAYIRKVDGPRSPSYGRSRSRSRSR

CDS

BASE COUNT 775 a 530 c 692 g 891 t
ORIGIN
Alignment Scores:
Pred. No.: 9,588-06 Length: 2878
Score: 95.00 Matches: 108
Percent Similarity: 56.46% Conservative: 10
Best Local Similarity: 51.67% Mismatches: 85
Query Match: 42.79% Indels: 6
DB: 10 Gaps: 2

US-10-014-927-19MOD_COPY_1_222 (1-222) x BC046773 (1-2878)

QY 9 IleTyrValGlyAsnLeuProGlyAspLeuArgLysCysGluValGluAspLeuPheTyr 28
Db 150 ATCTACGTGGGTACCTACCTCCGATATCCGACCAAGGACATCGAGACGTGTTTAC 209
QY 29 LysTyrGlyProIleValAspLeuAspLeuLysIleProProArgProProGlyTyrAla 48
Db 210 AAATACGGCGCATCCGACATCGACCTGAAGAACCCCGCGGGGACCGCCCTCCGCC 269
QY 49 PheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyr 68
Db 270 TTCGTGTAGTTCGAGACCCCGAGACCGGGAAGATCGGTGTACGTCGCGACGCTAC 329
QY 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArg-----Arg 85
Db 330 GACTACGAGGTACCGCTCGGTAGATTTCCCGAAGCGCGCGGGACCGCGCGA 389
QY 86 *****SerTyrSerAlaSerArgAlaProSerArg 105
Db 390 GCGCGCGCGGGGTGAGGCGCGCGCGCGAGAGCGCGCTATGGCGCGCGCTCCAGG 449
QY 106 ArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeu 125
Db 450 CGGTCCGGAACACAGATGGTGTCTCTGACTGCTCCGAGTGAAGTGGCAGACTTA 509
QY 126 LysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLys 145
Db 510 AAGGATCACATCGGTGAGGAGGTGATGTTATGTTACGTGATGTTTACCGAGAT----- 563
QY 146 GlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaLeuArgLys 165
Db 564 ---GGCACTGGTGTGAGTGTTCGAGGAGAAAGATATGACGTATGACGTTCGAGTT 620
QY 166 LeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyr 185
Db 621 CTGGATAACACTAAGTTAGATCTCAGCAGGAGAGAACTGCCCTACATCCGGGTTAAAGTT 680
QY 186 GluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArg 205

Db 681 GATGGCCGAGAGTCCAAAGTTATCGAAGATCTCGATCTCGAAGCCGTAGTCGTAGCAGA 740
QY 206 SerArgGlyProSerCysSerTyrSer 214
Db 741 ACCCGTAGCAGAGCAACAGCAGGAGT 767

Search completed: February 4, 2004, 23:06:47
Job time : 3518 secs